Biopolymer hairpin loops sustained by polarons

B. Chakrabarti,* B. M. A. G. Piette,† and W. J. Zakrzewski‡
Department of Mathematical Sciences, University of Durham, Durham DH1 3LE, United Kingdom
(Received 23 April 2012; revised manuscript received 11 July 2012; published 9 August 2012)

We show that polarons can sustain looplike configurations in flexible biopolymers and that the size of the loops depend on both the flexural rigidity of the polymer and the electron-phonon coupling constant. In particular we show that for single stranded DNA (ssDNA) and polyacetylene such loops can have as few as seven monomers. We also show that these configurations are very stable under thermal fluctuations and so could facilitate the formation of hairpin loops of ssDNA.

DOI: 10.1103/PhysRevE.86.021910 PACS number(s): 87.15.hm, 71.38.−k, 87.14.gk

I. INTRODUCTION

Conformational transitions of biopolymers as a result of the coupling between the electronic and elastic degrees of freedom are important for understanding native states of globular proteins and secondary structures of biopolymers such as DNA and RNA. In an attempt to understand toroidal states of DNA the globule-coil transition for semiflexible polymers in poor solvents has been explored using Brownian dynamics simulations [1,2]. The intermediate states arising in these systems have also been classified [1,2]. However, the collapse transition in polymers induced by polarons has been less explored [3]. Though there is a clear separation of energy scales between electronic phenomena occurring in the eV range and conformational elasticity of biopolymers operative in the k_BT range, a comparison of the individual contributions of these terms to the total energy indicates that their interplay can lead to novel phenomena hitherto unexplored.

Polarons are the result on the interaction between a free electron in the conducting band of a polymer chain and the phonons of that chain. They were discovered by Davydov [4–6], who proposed them as a mechanism to explain how energy can be transported along α helices in living cells.

In this paper we explore the possibility of polaron induced polymer-loop formation and stabilization arising in semiflexible chains. Our model is similar to the model proposed by Mingaleev *et al.* [3], who generalized the original model of Davydov [4,5] by incorporating long ranged electron-phonon interactions. In their work Mingaleev *et al.* showed that at zero temperature polarons can induce a spontaneous bend in a straight chain if the bending modulus is less than a critical threshold. Hence the bending can have many origins, and it is important to see whether its origin is electric in nature (i.e., induced by polarons) or due to some other (classical) interactions. A careful examination of the model, however, reveals that realistic polymers, e.g., DNA and polyacetylene, are more rigid, having their bending modulus 2 and 20 times above the threshold value, respectively. Therefore, although interesting from a theoretical point of view, the spontaneous bending on polymers induced by polarons is limited in scope when applied to physical systems. This does not imply that the polaron induced bending is irrelevant and so can be neglected. In fact it will always make a contribution, but for many physical systems this contribution may not be dominant. The folding free energies of single stranded DNA (ssDNA) hairpins have been measured in single molecule experiments using the optical traps [7,8]. The mechanism proposed here would alter these free energies, accounting for extra stability. In other words, the polaron is not a substitute for the binding between nucleic acids, but rather a stable state that can facilitate the formation of a hairpin when the nucleic acids are aligning optimally.

However, in a recent paper [9] we showed that the Mingaleev *et al.* model can explain spontaneous polaron transport on a chain having a bending gradient, e.g., α helices of light harvesting proteins. In this case, the bending of the chain is generated by the natural folding of the protein, which can induce a spontaneous polaron displacement. We showed that for a polymer configuration with a preferred bend, the polaron spontaneously accelerates along the bending gradient and gets reflected across sharply kinked junctions. Further we showed that at finite temperatures the polaron undergoes a biased random walk to a region of high curvature.

While polarons are not able to induce spontaneous conformational transitions in DNA and polyacetylene, because of their rigidity, they might sustain a folded configuration that might have been formed by other means, e.g., thermal fluctuations or mechanical stress. This is particularly true for ssDNA, whose bending modulus is only twice as large as the threshold value for spontaneous bending. This is what we are investigating in detail in this paper, which is organized as follows.

In Sec. II we review the Mingaleev *et al.* [3] model. In Sec. III we study loop configurations in which the last two nodes of the chains are held together by a polaron. We extend this analysis in Sec. IV to study hairpin-loop configurations for which the two opposite ends of the chain run parallel to each other, while the loop links the parallel strands together. Finally, we show that one can estimate analytically the value of the parameters for which loops can be formed in Sec. V. In Secs. VI and VII we look in some detail at loop and hairpin-loop configurations for both single stranded DNA and polyacetylene, and we show that the polarons in these two systems are very stable and that they can facilitate the formation of hairpin loops.

*buddhapriya.chakrabarti@durham.ac.uk
†b.m.a.g.piette@durham.ac.uk
‡w.j.zakrzewski@durham.ac.uk

II. MODEL

The model proposed by Mingaleev *et al.* [3] is described by the Hamiltonian

$$H = \sum_n \left[\frac{\hat{M}}{2}\left(\frac{d\vec{R}_n}{d\tau}\right)^2 + \hat{U}_n(\vec{R}) - \frac{1}{2}\Delta|\phi_n|^4 + W\left(2|\phi_n|^2 - \sum_{m\neq n} J_{nm}\phi_n^*\phi_m\right)\right], \quad (1)$$

where $\vec{R}_n$ describes the position of each chain node, $\hat{M}$ is the node mass, W is the linear excitation transfer energy, and Δ is the nonlinear self-trapping interaction. The excitation transfer coefficients $J_{n,m}$ are of the form

$$J_{n,m} = J(|\vec{R}_n - \vec{R}_m|) = (e^{\alpha} - 1)\, e^{-\alpha|\vec{R}_n - \vec{R}_m|/\hat{a}}, \quad (2)$$

where α^{-1} sets the relative length scale over which the interaction decreases, in units of $\hat{a}$, where $\hat{a}$ is the rest distance between two adjacent sites. The function $J_{n,m}$ describes the long range interaction between the electron field at different lattice sites n and m; its value decreases exponentially with the distance between them. Notice that when α is large and $|\vec{R}_n - \vec{R}_m| \approx \hat{a}$, this corresponds to a nearest neighbor interaction with $J_{n,m} \approx \delta_{n,m\pm1}[1 + \alpha(1 - \frac{|\vec{R}_n - \vec{R}_m|}{\hat{a}})]$.

In our formulation of the model, the normalization of the electron field is preserved, i.e.,

$$\sum_n |\phi_n|^2 = 1. \quad (3)$$

The phonon potential $\hat{U}_n$ consists of three terms:

$$\hat{U}_n(\vec{R}) = \frac{\hat{\sigma}}{2}(|\vec{R}_n - \vec{R}_{n-1}| - \hat{a})^2 + \frac{\hat{k}}{2}\frac{\theta_n^2}{[1 - (\theta_n/\theta_{\max})^2]} + \frac{\hat{\delta}}{2}\sum_{m\neq n}(\hat{d} - |\vec{R}_n - \vec{R}_m|)^2\Theta(\hat{d} - |\vec{R}_n - \vec{R}_m|), \quad (4)$$

where the Heaviside function is defined as $\Theta(x) = 1$ for $x > 1$ and $\Theta(x) = 0$ for $x < 1$.

The first two terms in $\hat{U}_n$ describe the elastic and the bending energy of the chain, respectively. $\hat{a}$ is the equilibrium separation between nodes, and θ_n is the angle between $\vec{R}_n - \vec{R}_{n-1}$ and $\vec{R}_{n+1} - \vec{R}_n$. Finally, $\theta_{\max}$ is the largest angle allowed between adjacent links.

The term proportional to $\hat{\delta}$ in $\hat{U}_n$ models hard-core repulsion between the atoms of the chain. $\hat{\delta}$ should always be larger than $\hat{\sigma}$, and $\hat{d}$ will correspond to the minimum distance allowed between nodes.

For convenience, the symbols denoted with a hat, e.g., $\hat{M}$, $\hat{\sigma}$, etc., correspond to physical variables carrying units and dimensions, while those without it correspond to dimensionless variables and parameters described below, except H, Δ, and W, which are dimensional quantities. We also use the symbol $\vec{R}$ for position of the nodes in physical units and $\vec{r}$ in dimensionless units. First, we define the time scale $\tau_0 = \hbar\Delta/W^2$ and use the lattice spacing $\hat{a}$ as the length scale. We can then define the dimensionless time t, position r, and coupling constant g as

$$t = \frac{\tau}{\tau_0}, \quad g = \frac{\Delta}{W}, \quad \vec{r} = \frac{\vec{R}}{\hat{a}}. \quad (5)$$

In terms of these variables the Hamiltonian takes the form

$$H = \frac{W^2}{\Delta}\sum_n \left[\frac{M}{2}\left(\frac{d\vec{r}_n}{dt}\right)^2 + U_n(\vec{r}) + g\left(2|\phi_n|^2 - \sum_{m\neq n} J_{nm}\phi_n^*\phi_m\right) - \frac{g^2}{2}|\phi_n|^4\right], \quad (6)$$

where

$$U_n(\vec{r}) = \frac{\sigma}{2}(|\vec{r}_n - \vec{r}_{n-1}| - a)^2 + \frac{k}{2}\frac{\theta_n^2}{[1 - (\theta_n/\theta_{\max})^2]} + \frac{\delta}{2}\sum_{m\neq n}(d - |\vec{r}_n - \vec{r}_m|)^2\Theta(d - |\vec{r}_n - \vec{r}_m|), \quad (7)$$

with

$$M = \hat{M}\frac{\hat{a}^2W^2}{\hbar^2\Delta}, \quad \sigma = \hat{\sigma}\frac{\hat{a}^2\Delta}{W^2}, \quad \delta = \hat{\delta}\frac{\hat{a}^2\Delta}{W^2}, \quad k = \hat{k}\frac{\Delta}{W^2}, \quad a = 1, \quad d = \frac{\hat{d}}{\hat{a}}. \quad (8)$$

We emphasize that we are dealing with a coarse-grained model of a polymer or biopolymer. Our model thus glosses over details such as the exact chemical nature of the molecules in question. However, as explained later, the parameter values used in our study correspond to synthetic and biological molecules and therefore should be amenable to detailed atomistic simulations and experiments.

Writing $\vec{r}_n = (x_{1,n}, x_{2,n}, x_{3,n},)$, we can derive the equation of motion for $x_{i,n}$ from the Hamiltonian (6):

$$M\frac{d^2x_{i,n}}{dt^2} + \Gamma\frac{dx_{i,n}}{dt} + F(t) + \sum_m \frac{dU_m}{dx_{i,n}} - g\sum_k\sum_{m<k}\frac{dJ_{km}}{dx_{i,n}}(\phi_k^*\phi_m + \phi_m^*\phi_k) = 0, \quad (9)$$

$$i\frac{d\phi_n}{dt} - 2\phi_n + \sum_{m\neq n} J_{nm}\phi_m + g|\phi_n|^2\phi_n = 0,$$

where the force $F(t)$ and the friction term $\Gamma\, dx_{i,n}/dt$ were added by hand to incorporate thermal fluctuations and $F(t)$ was chosen as a δ correlated white noise satisfying

$$\langle F(0)F(s)\rangle = 2\Gamma k_B T\delta(s), \quad (10)$$

where

$$k_B T = \hat{k}_B\hat{T}\frac{\hat{W}^2}{\hat{\Delta}} = \hat{k}_B\hat{T}\hat{W}g. \quad (11)$$

As the equation for x_i is expressed in units of $\hat{W}^2/(\hat{\Delta}\hat{a})$, we have $\Gamma = \hat{\Gamma}\hat{a}^2/\hbar$. The friction coefficient $\hat{\Gamma}$ can be evaluated from $\hat{\Gamma} \approx 6\pi\mu R_0$, where $\mu = 0.001$ Pa s for water and is up to 4 times that value for the cytoplasm and R_0 is the average radius of a single molecule of the lattice. Notice also that the electron field ϕ_n is coupled to the phonon field $x_{i,n}$ through the function J_{nm}.

In what follows we are primarily interested in stationary configurations. To compute such solutions numerically we chose an initial lattice configuration with a loop structure and localized the electron so that it overlapped with both tails of the loop. We achieved this by using an approximation for the polaron electron field and distributing it over a few nodes spread between the two ends of chain. In this way the polaron was able to bind the loop extremities together. We then relaxed the electron field, keeping the lattice configuration unchanged, and then evolved the entire system with an absorption term until it relaxed to a static configuration. This was achieved by solving Eq. (9) without thermal noise.

In all our simulations we started from a very small value of k, typically, $k = 0.005$, so that the lattice offered very little resistance to bending. We then increased the value of k in small increments using the relaxed conformation obtained for the previous k value as the initial configuration. We then equilibrated the system for the new value of k. By repeating the procedure for each value of g we have determined the critical value $k_{\rm crit}(g)$ up to which the given configuration can be sustained by the polaron.

Unless otherwise stated, we have used the following parameter values: $\delta = 10000$, $\sigma = 1000$, and $M = 0.5$. For stationary solutions the mass term does not affect the results, and δ was chosen so that the repulsion potential is close to that of a hard shell. Finally, for all the computed configurations, nothing prevents the nodes from being very close to their equilibrium distance, and hence we have selected a relatively large value for σ to approximate stiff cross-node links. Following Mingaleev *et al.*, we have also considered mainly the case $\alpha = 2$ and $d = 0.6$. Finally, we have also considered the effect of varying the values of these two parameters.

To solve Eq. (9) we used a fourth-order Runge-Kutta method with a time step $dt = 0.0001$ in dimensionless units. To compute static configurations, we took $T = 0$, i.e., no thermal noise, setting $\Gamma = 1$ and then integrating Eq. (9) until the system relaxed to a stationary solution.

To study the thermal stability of the configurations for DNA and polyacetylene, we solved Eq. (9), taking $T = 300$ K and estimated Γ from the radius of the molecules as described above. For those simulations we started from the static configuration for which we wanted to evaluate the stability and let the system thermalize itself. The time needed for this thermalization was always orders of magnitude smaller than the average lifetime of the configurations we considered, and so we did not need to resort to a sophisticated thermalization procedure as we did in [9].

III. PLAIN-LOOP CONFIGURATIONS

Our first investigation involved considering simple-loop configurations for which all the nodes lie more or less on a circle with the two end points close to each other (separated by a distance d). When k is very small, the favored configuration is one similar to the one presented in Fig. 1(a). In Fig. 1(a), the electron probability density is represented by the color of the node. A dark color corresponds to a null value, while a light color corresponds to a higher probability density. The node at which the polaron field has its maximum value is close to the last two points at the opposite ends of the chain. This allows the electron field to be distributed on two nearby nodes rather than a single one, and as k is small, the deformation of the chain does not prevent this from taking place. As k increases, such a localization becomes energetically expensive, and the configuration assumes the shape of a horseshoe, as presented in Figs. 1(b) and 1(c). As one increases k further, there is a point at which the stretching energy is too large, and the polaron is not able to sustain the loop anymore.

The difference between Figs. 1(b) and 1(c) is that in the former the electron is localized equally on the two end nodes, while in the latter it is localized mostly on a single node. The difference is dictated by the value of g: for small g, the polaron is wide, and the electron spreads itself nearly equally between the two end points of the chain [Fig 1(b)]. As g increases, the polaron becomes more localized and the electron becomes localized, more asymmetrically, on a single node [Fig. 1(c)].

The critical value of k as a function of g is presented in Fig. 2(a) for loops consisting of 9 to 14 nodes. It is interesting to note that when g is small, the critical value of k is small. This can be explained by the fact that the coupling parameter g is small and also by the fact that the polaron is delocalized and hence the fraction of the electron close to the end point is smaller than for larger values of g. The maximum value of $k_{\rm crit}$ is reached for $g \approx 5$. For very large values of g, the electron is nearly fully localized on a single lattice point, but the attraction exerted by the polaron surprisingly decreases very slowly.

Having followed [3] and taken the values $\alpha = 2$ and $d = 0.6$ for the results presented so far, it is worth checking how these two parameters affect the results that we have obtained. We started by varying α, which controls, through $J_{n,m}$, how fast the coupling between nodes decreases with the distance separating them. The results are presented in Fig. 2(b), where we see that $k_{\rm crit}$, contrary to what one might expect, increases with α. This is easily explained: having chosen $d = 0.6$, increasing α not only reduces the long distance interaction between nodes but

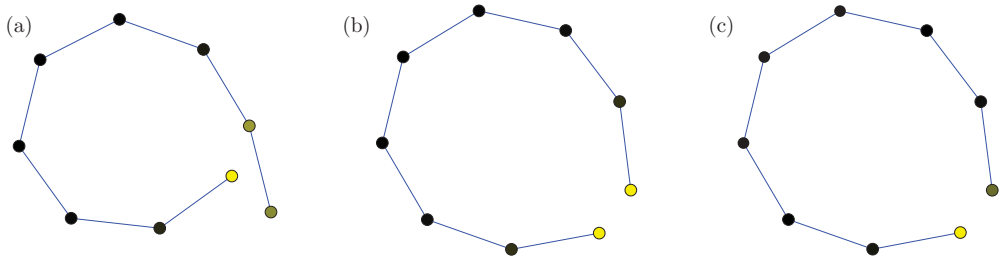

FIG. 1. (Color online) Loop configuration for $N = 9$ nodes for (a) $g = 2.5$, $k = 0.5$, $|\phi_0|^2 = 0.407$, $|\phi_8|^2 = 0.221$, $|\phi_7|^2 = 0.249$, (b) $g = 2.5$ $k = 4$, $|\phi_0|^2 = |\phi_8|^2 = 0.397$, and (c) $g = 5$ $k = 5$, $|\phi_0|^2 = 0.642$, $|\phi_8|^2 = 0.272$.

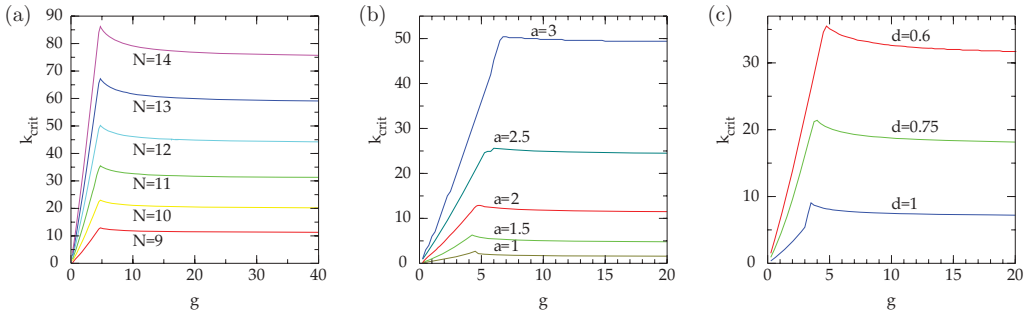

FIG. 2. (Color online) Critical value of k for the existence of a loop configuration. (a) $\alpha = 2$ and $N = 9$ to 14 nodes. (b) $N = 9$ nodes, $\alpha = 1,1.5,2,25,3$, and $d = 0.6$. (c) $N = 11$ nodes, $d = 0.6,0.75,1$, and $\alpha = 2$.

also increases exponentially the binding energy of nodes that are very close to each other. The binding energy of the end nodes, which are separated by a distance $d < 1$, thus increases with α. For this reason, we have decided to take $d = a = 1$ when we consider single stranded DNA and polyacetylene later in the paper.

In Fig. 2(c) we show how the critical value of k varies with d. As the parameter d sets the minimal distance allowed between two nodes and given that $J_{n,m}$ decreases with the distance between nodes, it is not surprising that k_{crit} decreases when d becomes larger, but loop configurations can still be held by the polaron.

IV. HAIRPIN-LOOP CONFIGURATIONS

Now we consider a hairpin-loop configuration as presented in Fig. 3, which is similar to the structure that single stranded DNA can form and which is potentially more relevant to long chains. As for the plain loops, we generated these configurations for a small k and then slowly increased its value until the number of links L making the loop increased by one unit. This gave us the critical value k_{crit} for which the hairpin-loop configuration of a given size can be sustained by the polaron.

The results are presented in Fig. 4, where we can see a sharp transition around $g = 10$. Below this value the hairpin loops are only viable for relatively small values of k, but above it, they are sustainable for much more rigid chains. This is due to the fact that for small values of g, the polaron is always distributed over the handle of the hairpin loop, while when $g > 10$, it is localized mostly on one lattice site, at the base of the loop. In that case the interaction is stronger and supports loops for larger values of k.

To make sure this was not an artifact of our procedure, we have tried to construct solutions using various initial conditions. We also used solutions obtained for $g > 10$ as initial conditions and then slowly decreased the value of g. Regardless of the procedure we used, we always obtained the curve of Fig. 3(a). As expected, the configurations of Fig. 3 are harder to sustain than a simple loop as the chain needs to be bent near the stem of the hairpin loop.

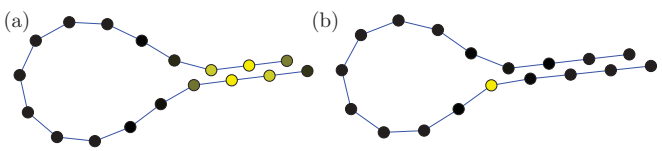

FIG. 3. (Color online) Hairpin-loop configuration for $N = 18$ nodes. The brightness of the nodes is proportional to $|\phi|^2$. (a) $g = 1.5$ and $k = 0.5$ (max $|\phi|^2 = 0.188$) and (b) $g = 11$ and $k = 10$ (max $|\phi|^2 = 0.887$).

V. ANALYTIC APPROXIMATION

Having computed numerically the critical value $k_{\text{crit}}(g)$ for which the polaron is able to sustain a loop of a given size, we now try to estimate this value analytically. To do this, we consider a circular configuration of radius R made out of N segments, with one segment of length b and $N - 1$ segments of length a, as depicted in Fig. 5. Note that the two nodes separated by the distance b are not linked to each other. If ξ and μ are the angles opposite a and b, respectively, we have

$$(n-1)\xi + \mu = 2\pi, \quad \sin\left(\frac{\xi}{2}\right) = \frac{a}{2R}, \quad \sin\left(\frac{\mu}{2}\right) = \frac{b}{2R}, \tag{12}$$

so

$$\sin\left(\frac{\xi}{2}\right) = \frac{a}{b}\sin\left(\frac{\mu}{2}\right) = \frac{a}{b}\sin\left(\frac{(n-1)\xi}{2}\right). \tag{13}$$

Choosing specific values $a = 1, b = b_0$, this transcendental equation can be solved numerically to obtain the corresponding value $\xi = \xi_0$. We can then perform a first order expansions

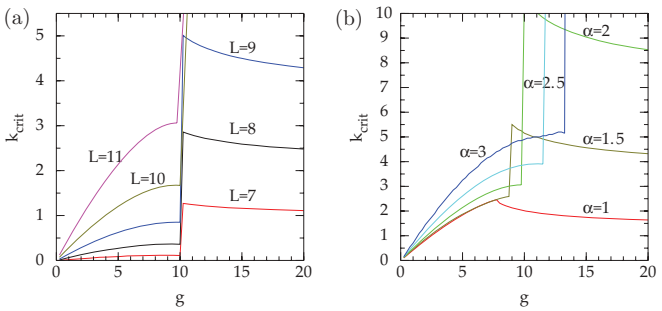

FIG. 4. (Color online) Critical value of k for the existence of a hairpin-loop configuration. (a) $\alpha = 2$ and $N = 7$ to 11 node loops. (b) $N = 9$ nodes and $\alpha = 1,1.5,2,2.5,3$.

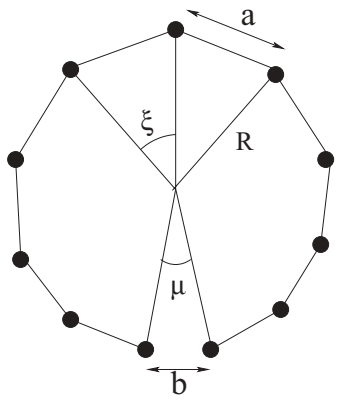

FIG. 5. Schematic representation of a polymer loop with N nodes. $N-1$ bonds with rest distance a subtend an angle ξ at the center. The two end nodes, on which the electron is localized, are separated by a distance b and span an angle μ at the center.

around this solution, $b = b_0 + \delta b, \xi = \xi_0 + \delta\xi$, and we obtain

$$\delta\xi = \frac{\sin\left(\frac{\xi}{2}\right)}{\frac{N-1}{4}\cos\left[(N-1)\frac{\xi}{2}\right] - \frac{b_0}{4}\cos\left(\frac{\xi}{2}\right)}\delta b. \tag{14}$$

To determine the critical values of g and k for which a loop configuration can exist, we have to minimize the Hamiltonian and, for each value of g and b, determine the value of k for which this Hamiltonian has a minimum. For each g, we then select the value of b for which k is the largest.

Let us assume that the loop is symmetric, so that the $N-2$ bending terms are all identical and are functions of ξ. The elastic terms are then also equal, but as they do not depend on ξ, they are constant and can thus be ignored for the minimization. The repulsion term proportional to δ can also be ignored if $b > d$. When $b < d$, the repulsion term leads to a very large energy increase, and we can thus consider that $b = d$ is the smallest value we should consider.

To evaluate the electron field, we take the continuum limit of Eq. (9) for stationary solutions:

$$\frac{d^2\phi_c}{dx^2} + \hat{g}|\phi_c|^2\phi_c - \lambda\phi_c = 0, \tag{15}$$

which is the well-known nonlinear Schrödinger equation, where $\hat{g} = g/(1-e^{-\alpha})$ and which admits the following solution:

$$\phi_c(x) = \sqrt{\frac{\hat{g}}{8}}\cosh\left(\frac{\hat{g}x}{4}\right). \tag{16}$$

Note that $\int |\phi(x)|^2 dx = 1$ and $\lambda = -\hat{g}^2/16$. From the numerical solutions, we know that the wave function is centered on one of the two end lattice points, and we can thus take $\phi_0 = \phi_c(0)$ and $\phi_{N-1} = \phi_c(b_0)$.

The Hamiltonian can then be approximated by the following function of b_0 and ξ_0:

$$H \approx -2\,g\,(e^{\alpha}-1)e^{-\alpha b_0/a}\phi_0\phi_{N-1} + \frac{k}{2}\frac{(N-2)\xi_0^2}{\left[1-\left(\frac{\xi_0}{\theta_{\max}}\right)^2\right]}. \tag{17}$$

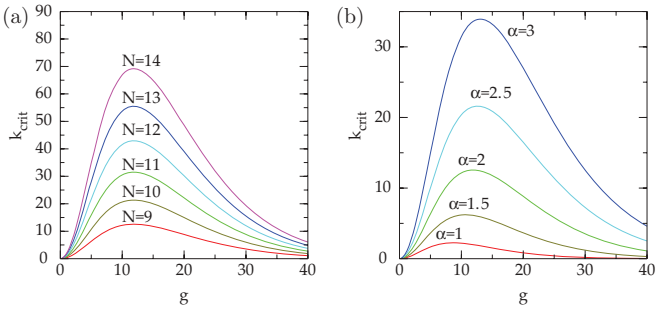

FIG. 6. (Color online) Theoretical estimation of the critical value of k for the existence of a loop configuration as a function of α and N. (a) $\alpha = 2$ and $N = 9$ to 14 nodes. (b) $N = 9$ nodes and $\alpha = 1, 1.5, 2, 25, 3$.

Next we compute the variation of H with respect to b and ξ:

$$\delta H = 2g\alpha\phi_0\phi_{N-1}(e^{\alpha}-1)e^{-\alpha b_0}\delta b + \frac{k(N-2)\xi_0}{\left[1-\left(\frac{\xi_0}{\theta_{\max}}\right)^2\right]}\left[1-\xi_0^2\left(1-\frac{1}{\theta_{\max}^2}\right)\right]\delta\xi. \tag{18}$$

Using Eq. (14) and imposing $\delta H = 0$, we get the condition

$$k_{\rm crit} = \frac{g}{D}\phi_0\phi_{N-1}, \tag{19}$$

where

$$D = \frac{1}{2\alpha(e^{\alpha}-1)e^{-\alpha b_0}}\frac{(N-2)\xi_0}{\left[1-\left(\frac{\xi_0}{\theta_{\max}}\right)^2\right]}\left[1-\xi_0^2\left(1-\frac{1}{\theta_{\max}^2}\right)\right] \times \frac{-4\sin\left(\frac{\xi_0}{2}\right)}{(N-1)\cos\left[(N-1)\frac{\xi_0}{2}\right] - b_0\cos\left(\frac{\xi_0}{2}\right)}, \tag{20}$$

which depends on g but not on k. From Eqs. (19) and (20) it is clear that $k_{\rm crit}$ increases as b_0 decreases, and so we have to choose the smallest possible value for b_0, i.e., $b_0 = d$.

Taking $a = 1$, $b_0 = d = 0.6$, we can solve Eq. (13) to obtain the values of ξ_0 listed in Table I, which we can use to estimate $k_{\rm crit}$. The results are presented in Fig. 6, from which we see that our evaluation reproduces the gross features of the results obtained numerically (Fig. 4): $k_{\rm crit}$ is small when g is very small and increases with g until a maximum is reached. Then, as g increases further, $k_{\rm crit}$ slowly decreases. The maximum value obtained for $k_{\rm crit}$ is slightly smaller than the numerical value obtained before. The biggest discrepancy between the numerical and analytic results is for large g, but this is to be expected as this is the limit where the polaron is strongly localized and so is less well approximated by Eq. (16).

VI. SINGLE STRANDED DNA

Having considered the Mingaleev *et al.* polaron model in general, we now consider two explicit cases: ssDNA and polyacetylene, both of which have parameters allowing the polaron to sustain loops. Hairpins structures and cruciforms of ssDNA, as well as the similar RNA, play an important role in biological processes such as regulation of transcription. [10–12].

TABLE I. Value of ξ_0 for various numbers of nodes N.

	N					
	9	10	11	12	13	14
ξ_0	0.731341	0.654972	0.593071	0.541875	0.498824	0.462115

The parameters of our model were obtained from several sources. First of all, $\hat{k}$ can be determined from the flexural rigidity, $\hat{k} = \lambda \hat{k}_B \hat{T}/R_0$, where R_0 is the radius of the DNA strand. We do not have experimental values of $\hat{\delta}$, but its actual value does not play an important role except that it must be large enough to mimic a hard-shell repulsion. In practice, we chose a value larger than $\hat{\sigma}$.

For single stranded DNA we have $R_0 \approx 0.33$ nm [13], $\Delta \approx 0.4$ eV, $W \approx 0.3$ eV [3], $\hat{k} \approx 0.11$ eV [13], and $\hat{\sigma} \approx 1.5\,\text{eV/A}^2$ [14]. This leads to the following dimensionless values:

$$g \approx 1.33, \quad \sigma \approx 72.51, \quad k \approx 0.487, \quad M \approx 2.5 \times 10^5,$$

$$k_B T = k_B \hat{T} \frac{\hat{\Delta}}{\hat{W}^2} \approx 0.115, \quad \Gamma \approx 3210, \quad \tau_0 \approx 2.92 \times 10^{-15}\ \text{s}$$

We thus see that single stranded DNA sits at the bottom left region of Figs. 2, 4, and 6. For our simulations, we have chosen $\alpha = 2$, $\delta = 100$, and $d = 1$; the latter parameter was taken as the worst case we could consider. We then found that DNA can easily sustain loops of 10 segments and hairpin loops with 11 segments. We then studied the thermal stability of the configurations that we have obtained at $T = 300$ K. To achieve this, we started from a static configuration that we had obtained for DNA. The energies of the loop configurations were as follows: $E = -0.04$ eV ($N = 9$), $E = -0.225$ eV ($N = 10$), and $E = -0.32$ eV ($N = 11$). We thus see that the binding energy of the $N = 9$ loops has a binding energy of the order of $k_B T$, and we thus expect it to be relatively unstable. For $N = 10$ and 11, the binding energy is larger, and we thus expect these configurations to be much more stable.

We then solved Eq. (9), including the thermal noise, and ran 100 simulations for an extended period of time. We started by running 100 simulations for a loop made out of 10 nodes, and we measured an average unfolding time of 1 ns. We also observed that a loop made out of 11 nodes is much more stable and experiences a very slow unfolding of the loop with an average decay time of approximately 1 μs, a time scale that is relatively long from a biomolecular point of view.

We have also performed thermal simulations for a hairpin-loop configuration of 18 nodes and $L = 11$ and did not observe a single unfolding of the chain in 20 μs. As the integration time step was approximately 3×10^{-17} s this required 100 simultaneous simulations each performing around 10^{10} integration steps, and we decided not to run it any longer as the stability of the loop was sufficiently well established.

Under thermal noise, the stem, made out of the two parallel ends of the chain, deforms itself, and the chain takes the shape of a loop where the polaron links the two opposite ends of the chain around a couple of nodes, as presented in Fig. 7. In our simulations we observed that as the polaron moves along the chain, the size of the loop that it formed fluctuated constantly in time but it never unfolded. We can thus conclude that the DNA polarons loops are very stable.

This loop configuration could play an important role in the formation of single stranded DNA hairpin loops *in vivo*. The formation of such configurations depends on the likelihood of complementary DNA bases to face each other before they can bind by hydrogen bonding, and this likelihood decreases rapidly as the length of the chain increases [15] and partially matching DNA base sequences are less stable than perfectly matched ones [16]. While homogeneous sequences of DNA bases can bind quite rapidly, such as the one used in [15] and [16], for example, irregular sequences, such as ATGCAGTC...GACTGCAT are less likely to match purely randomly. As the polaron folds the ssDNA into a loop and moves along the chain, the DNA bases on the opposite segments of the loop slide relative to each other, under the action of thermal excitation, increasing the probability for a complementary sequence of bases to face each other and bind. *In vitro*, the reaction time of hairpin loops has been determined to be several microseconds [15], a length of time that, as we have shown, a DNA polaron can outlive easily. Hence, we can conclude that DNA polarons can increase the rate of formation of hairpin loops.

In our model, we have not taken into account the effect of water on the polaron. Recent studies [17,18] have suggested that its effect would be to reduce the polaron size, which, as can be seen from Eq. (16), corresponds to increasing the value of g. The net effect would thus be to move DNA to a parameter region where the polaron is stronger, as can be seen in Figs. 2 and 4. As a result the polaron would then be able to sustain smaller loops.

VII. POLYACETYLENE

For polyacetylene, the physical value of the parameters are given by $R_0 \approx 0.24$ nm, $W \approx 2.5$ eV, $\Delta \approx 6.4$ eV,

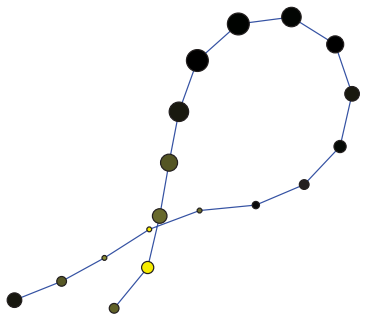

FIG. 7. (Color online) Thermalized, $T = 300$ K, hairpin-loop DNA configuration for $N = 18$ nodes. The sizes of the disks are an exaggerated indication of their depth in the direction transverse to the plane of view. (The two nodes close to the crossing point are separated by the same distance as two neighbor points.) The brightness of the nodes is proportional to $|\phi|^2$.

$\hat{\sigma} \approx 21$ eV/A^2 [19], and $\hat{k} \approx 3.7$ eV [20], and so

$$g \approx 2.56, \quad \sigma \approx 128, \quad k \approx 3.79, \quad M \approx 3.2 \times 10^4,$$
$$k_B T \approx 0.026, \quad \Gamma \approx 1316, \quad \tau_0 \approx 6.74 \times 10^{-16} \text{ s}.$$

Once again, we took $d = 1$ and $\alpha = 2$ to avoid the potentially spurious effects induced at close distance by $J_{n,m}$. In this case, we were able to make loops out of 12 nodes and hairpin loops of 12 nodes too.

Under thermal fluctuation, both were very stable. In this case, the integration time step was approximately 7×10^{-18} s, and our attempt to evaluate the configuration average lifetime was achieved by running 100 simultaneous each performing over 10^{10} integration steps; we did not observe a single unfolding of the chain in 13 μs. We also ran simulations for a hairpin-loop configuration of $N = 18$ nodes and $L = 12$ and also did not observe a single unfolding of the chain in over 10 μs. This is not surprising as the energy of that configuration was $E = -2.53$ eV.

VIII. CONCLUSIONS

In this paper we have studied the possibility of a polaron to sustain loops and hairpin-loop configurations. In these configurations the polaron was localized over lattice nodes that were well separated along the chain backbone but spatially close to each other because of the bending of the chain. The polaron then acted as a linker between the two regions of the chain and so could sustain the loop configuration if the chain was not too rigid. The Mingaleev *et al*. model we have used to describe this property takes into account the long distance interactions between the electron and the phonon field, with a strength decreasing with the distance. For the configurations we have studied, the most important contribution comes from lattice nodes that are spatially close to each other, and the energy contribution from the next to nearest neighbor is not essential, unlike in our study of spontaneous polaron displacements [9], where the next to nearest neighbor terms were essential for the polaron to move along the bending gradient of the chain.

We have determined the critical value of the chain rigidity k_{crit} as a function of the polaron coupling constant g, and we have shown that polarons are able to sustain relatively small loops for a wide range of parameters values. We have then shown that both DNA and polyacetylene are flexible enough for a polaron to sustain hairpin-loop configurations. Moreover, we have also shown that these hairpin-loop configurations are very stable under thermal excitations, with average lifetimes exceeding 10 μs, and that they can facilitate the formation of hairpin loops of single stranded DNA. At present we are not aware of any experimental work in this direction and hope that our work will spark some interest among experimentalists. We believe that polaron formation and polaronic mode of transport can be picked up by Raman spectroscopic studies. Further, we hope that quantum mechanical calculations of hairpin formation incorporating polarons might help narrow down the parameter range where such phenomena might be observed.

ACKNOWLEDGMENT

B.C. was partially supported by EPSRC Grant No. EP/I013377/1.

[1] B. Schnurr, F. C. Mackintosh, and D. R. Williams, Europhys. Lett. **51**, 279 (2000).
[2] B. Schnurr, F. Gittes, and F. C. MacKintosh, Phys. Rev. E **65**, 061904 (2002).
[3] S. F. Mingaleev, Y. B. Gaididei, P. L. Christiansen, and Y. S. Kivshar, Europhys. Lett. **59**, 403 (2002).
[4] A. S. Davydov, J. Theor. Biol. **38**, 559 (1973).
[5] A. S. Davydov, Phys. Scr. **20**, 387 (1979).
[6] A. S. Davydov, Phys. D **3**, 1 (1981).
[7] M. T. Woodside, W. M. Behnke-Parks, K. Larizadeh, K. Travers, D. Herschlag, and S. M. Block, Proc. Natl. Acad. Sci. USA **103**, 6190 (2006).
[8] A. N. Gupta, A. Vincent, K. Neupane, H. Yu, F. Wang, and M. T. Woodside, Nat. Phys. **7**, 631 (2011).
[9] B. Chakrabarti, B. M. A. G. Piette, and W. J. Zakrzewski, Europhys. Lett. **97**, 4705 (2012).
[10] R. M. Wadkins, Curr. Med. Chem. **7**, 1 (2000) [http://www.ingentaconnect.com/content/ben/cmc/2000/00000007/00000001/art00002].
[11] A. M. Gacy, G. Goellner, N. Juranic, S. Macura, and C. T. McMurray, Cell **81**, 533 (1995).
[12] M. A. Glucksmann-Kuis, X. Dai, P. Markiewicz, and L. B Rothman-Denes, Cell **84**, 147 (1996).
[13] M. Mandelkern, J. Elias, D. Eden, and D. Crothers, J. Mol. Biol. **152**, 153 (1981).
[14] S. B. Smith, Y. Cui, and C. Bustamante, Science **271**, 795 (1996).
[15] G. Bonnet, O. Krichevsky, and A. Libchaber, Proc. Natl. Acad. Sci. USA **95**, 8602 (1998).
[16] M. Kenward and K. D. Dorfman, J. Chem. Phys. **130**, 095101 (2009).
[17] S. M. Kravec, C. D. Kinz-Thompson, and E. M. Conwell, J. Phys. Chem. B **115**, 6166 (2011).
[18] A. K. Thazhathveetil, A. Trifonov, M. R. Wasielewski, and F. D. Lewis, J. Am. Chem. Soc. **133**, 11485 (2011).
[19] A. J. Heeger, S. Kivelson, J. R. Schrieffer, and W.-P. Su, Rev. Mod. Phys. **60**, 781 (1988).
[20] F. L. VanNice, F. S. Bates, G. L. Baker, P. J. Carroll, and G. D. Patterson, Macromolecules **17**, 2626 (1984).